Szybciej nie zawsze znaczy lepiej

Przykład: szkic odpowiedzi o status zamówienia. Człowiek sprawdza treść i decyduje o wysłaniu. Próba obejmuje standardowe, zmienione i złożone sprawy.

CZAS BEZ AI

4.92 h

Te same 12 zadań

CZAS Z AI

2.40 h

Wraz z kontrolą i poprawkami

BILANS PO URUCHOMIENIU

1.02 h

Nakład uruchomienia: 1,50 h

JAKOŚĆ

10 / 12

83,3% przy wymaganiu 95%

Decyzja: poprawić instrukcję i ponowić ocenę. Dodatni bilans czasu nie wystarcza. Dwie złożone sprawy nie spełniły wzorca odpowiedzi; trzeba wyjaśnić źródło błędu.

Co zapisujesz	Po co
Przypadek i typ	Nie ukrywaj trudnych spraw w średniej ze standardowych.
Czas porównawczy	To samo zadanie, podobna trudność i wymagany wynik.
Pełny czas z AI	Przygotowanie + generowanie + kontrola + poprawki.
Jakość	Spójny wzorzec oceny: fakty, kompletność, odbiorca, następny krok.
Błąd krytyczny	Oddzielna flaga i przegląd niezależnie od średniego wyniku.
Wersja instrukcji	Pozwala powtórzyć próbę i porównać zmianę.

Przykład nie prognozuje oszczędności w Twojej firmie. Licz pełną pracę, także po stronie osób kontrolujących.

Instrukcja, wyjątki i akceptacja

Dobra instrukcja zaczyna się od wzorca pracy. Agent dodatkowo wykonuje kroki i używa narzędzi, dlatego potrzebuje jasno określonych uprawnień.

Zadanie: przygotuj szkic odpowiedzi na pytanie o status zamówienia.
Wejście: pytanie, zweryfikowany status i obowiązujący standard.
Wynik: krótkie podsumowanie, szkic odpowiedzi, źródła faktów i lista braków.
Zasady: nie wymyślaj terminów ani przyczyn. Przy brakujących lub sprzecznych danych zatrzymaj opracowanie faktów i wskaż, kto ma je potwierdzić.
Kontrola: zgodny odbiorca i status, brak nieuprawnionych obietnic, jasny następny krok.

Krok agenta	Granica i dowód
1. Odczyt	Tylko zatwierdzone źródła; zapisz wersję i czas danych.
2. Opracowanie	Szkic oraz wskazanie braków; bez domyślania faktów.
3. Ocena	Wzorzec jakości i jawne wyjątki; błąd krytyczny zatrzymuje próbę.
4. Akceptacja	Wysyłka i zmiana danych wymagają decyzji uprawnionej osoby.
5. Rejestr	ID sprawy, źródła, wersja instrukcji, wynik oceny i decyzja.

Ćwiczenie: wybierz 12 zróżnicowanych, dopuszczonych do użycia przypadków. Zapisz wzorzec oceny przed testem. Oceń oba sposoby pracy, przeanalizuj każdy błąd i sprawdź poprawioną instrukcję także na przypadkach wcześniej poprawnych.

Podstawa zarządzania ryzykiem: NIST AI RMF, funkcje Govern, Map, Measure i Manage.
airc.nist.gov/airmf-resources/airmf/5-sec-core/